USABILITY EVALUATION OF A VIRTUAL LEARNING ENVIRONMENT: A UNIVERSITY CASE STUDY

Attila Vertesi[1], Huseyin Dogan[2], Angelos Stefanidis[2], Giles Ashton[2] and Wendy Drake[2]
[1]*Bournemouth & Poole College, North Rd. Poole, BH14 0LS, United Kingdom*
[2]*Bournemouth University, Talbot Campus, Fern Barrow, Poole, BH12 5BB, United Kingdom*

ABSTRACT

As part of educational technology developments in Higher Education, every university has adopted a Virtual Learning Environment (VLE) that facilitates online methods of delivery by enabling the submission of course materials, course management system and computer-mediated communication. VLE is regarded as technology which is either accepted or rejected by its users such as students, academics and administrators. Perceived usefulness and ease of use play an important role in user acceptance and satisfaction. This paper provides quantitative results of usability evaluations i.e., the System Usability Scale (SUS) scores from different user groups including students (n=137), academics (n=23), administrators (n=19) and learning technologist (n=3). The qualitative element of the VLE evaluation comprised the utilization of an approach called Interactive Management (IM) (n=13). The results showed that the newly implemented VLE performed under the average usability expectation (SUS score of 58.6). Students on average evaluated the usability of the VLE higher than the staff. The usability scores of the students from different courses showed remarkable differences. The ranked and categorised feedback from the IM session highlights the importance of planning, training and communication before and during the implementation, as well as the aspect of usability and learnability of the VLE.

KEYWORDS

Virtual Learning Environment (VLE), Learning Management System (LMS), Usability Evaluation, System Usability Scale (SUS), Interactive Management (IM), Technology Acceptance Model (TAM)

1. INTRODUCTION

Information technology is an essential component of the educational technology in Higher Education. Virtual Learning Environment (VLE) and Learning Management System (LMS) are often used as synonyms (Paulsen, 2002, p.6) describing a complex information technology system that integrates course management tools for the course administrators, online accessibility of learning materials and assignments; as well as a communication and collaboration platform for the students and lecturers (Ryan et al, 2013). The quality and usability of a VLE are the key features for its success by influencing user satisfaction and acceptance (Babić, 2012). Usability is the extent to which users can use a product or service to achieve specified goals efficiently and effectively while promoting feelings of satisfaction in a given context of use (ISO 9241-11). There are two aspects of usability in educational technology: technical usability and pedagogical usability (Melis et al, 2003). Technical usability refers to the Human-Computer Interaction (HCI), while pedagogical usability aims at supporting the learning process. Perceived usefulness and perceived ease of use represent usability as main factors in the Technology Acceptance Model (TAM) (Davis et al, 1989). TAM tries to predict the acceptability of a new technology.

The university has dedicated VLEs have been in operation for more than 12 years, currently used by over 20,000 students and 2,000 staff. Under EU regulations, the university was required to go out to tender for a new VLE at the end of the contract with the current VLE supplier. In total, 250 students and staff members, representing ten departments participated in the selection of the new VLE. More than 50 staff worked on the preparation and implementation of the new VLE for 8 months before it was introduced in September 2017 in a phased implementation. During phase 1, 40% of the students were transferred to the new system. The university moves towards, the full rollout from September 2018. The new VLE offers a personalised learning experience with learning analytics capabilities, integrated social media, chat, video features and game-based

learning for the students. Staff can benefit from the customisable course development, program management, account management, training, and end-user help desk support. In order to support the decision making for selecting the most appropriate VLE for the university, a preliminary usability evaluation has been carried out in 2016 on the three VLEs remaining in contention during the final stages of a rigorous procurement process.

In this research, a follow-up usability examination has been carried out on the implemented new VLE utilising the same methodology (SUS) in April 2018, six months after its launch. The result contributes to the next phase of the implementation process and provides feedback to the implementation team.

System Usability Scale (SUS) (Brooke, 1996) was utilised to carry out a general quantitative usability evaluation. The SUS scores from different user groups were analysed and compared. More detailed, factor analysis was applied where the low usability scores demanded it.

Interactive Management (IM) (Broome and Keever, 1986) methodology was applied to facilitate effective group communication to receive detailed feedback about the usability and the implementation of the new VLE.

The aim of this paper is twofold: (a) to provide reliable quantitative and qualitative feedback of the usability and the implementation of the new VLE (b) to offer a proposal on how to apply different evaluation methods to assess the usability of an educational technology system, e.g., VLE by the combined application of SUS and IM methodology.

In section 2, related usability studies are discussed in, especially SUS application on VLEs. Section 3 introduces the methodologies used, SUS and IM. The results are discussed and interpreted in section 4 and conclusions presented in section 5.

2. USABILITY STUDIES ON VIRTUAL LEARNING ENVIRONMENTS

A growing number of studies examine the usability of the VLE by utilizing SUS as a methodology. In 2006, a web-based e-learning platform called SPIRAL was developed and evaluated (Renaut et al., 2006) at University Claude Bernard Lyon 1. Although the SUS ratings have not been published, 72% of the professors found the system usable, according to the paper.

Three different e-learning platforms were measured using SUS by Ayad and Rigas (2010). User performance, learning effectiveness and satisfaction were examined to explore the usability aspects of the system. The three platforms were Virtual Classroom, Game-based and Storytelling. The SUS scores for the three platforms were 75.3, 73.4 and 64.5 respectively. The Storytelling scored a little behind the other two.

An interesting comparative research article was published (Gallud et al., 2013) regarding the usability enhancement of the Moodle LMS. The study examined the performance of the system in remote collaboration. The SUS score of the original Moodle system in these features was 46.75, which indicates serious usability problems. Using a different collaborative tool called Drag&Share within Moodle, the usability of the LMS enhanced dramatically. The SUS score increased significantly to 89.5 after the implementation of Drag&Share, which indicates a very good usability in the remote collaboration feature.

There is a very rare longitudinal study about a simulation-based learning system (Luo et al., 2014), that measured the perceived usability of the students after the first semester and after the second semester. Initially, the SUS score was 58.1, suggesting that the system needed improvement. Based on the collected data, the system had been modified, and after the second semester, the score rose to 65.9. Following another development for teachers, they evaluated the new module to 74.45, showing their satisfaction. This research also highlights the perceived usability of different user groups (e.g., teachers and students) may vary.

The above-mentioned divergence between the perceived usability of students and teachers is discussed by Emelyanova and Voronina (2014). The various aspects of the VLE and the difference between the perception of the usability should be considered when making a decision about the improvement of the system.

A comprehensive usability study was conducted in nine European secondary schools, all using UNITE e-learning platform, with the participation of 23 teachers and 47 students (Granic and Cukusic, 2011). Teachers evaluated the system at 53.15 and students gave 59.36 in average using the SUS questionnaire. The difference between the perception of the usability is also noticeable in this study. However, in this case, the students scored the system higher than the teachers.

3. METHODOLOGY

3.1 System Usability Scale (SUS)

Usability evaluation consists of methodologies for measuring the usability aspects of a system's user interface and identifying specific problems (Nielsen, 1993). There are numerous methods available for assessing the usability of a product (Brooke, 1996; Lewis, 1991, 1995; Tullis and Albert, 2008). System Usability Scale (SUS) is one of the most widely adopted methods (Brooke, 1996) due to its shortness, simplicity, comprehensiveness and reliability even with a small sample size (Tullis and Stetson, 2004).

SUS, developed by Brooke (1996), is a 10-item scale (Fig. 1). The ten statements can be rated on a five-point (Likert-type) scale ranging from 'strongly disagree' to 'strongly agree'. The SUS score, calculated from the answers, is a number between 1 and 100 which can easily be compared to the SUS scores of other similar or different systems and products. This methodology provides a reliable quantitative result of the usability of the VLE (Orfanou et al., 2015) that can highlight potential usability issues but does not identify them or give in-depth analysis about the possible causes.

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.

Figure 1. SUS questionnaire

3.2 Interactive Management (IM)

Interactive Management (IM) is a methodology designed to manage complex or new organisational or technical problems associated with multiple disciplines, involving different departments (Broome and Keever, 1986). IM offers methods to facilitate effective communications, promotes consensus-based decision making through idea generation, structuring and design. IM methods can be used to gather the requirements, needs, demands and ideas of the stakeholders for a better understanding of the problem space (Dogan and Henshaw 2010). IM tools are utilised to obtain feedback from the users about the implementation of the new VLE system. IM involves three phases: Planning, Workshop and Followup (Warfield and Cárdenas 1994). During the workshop, Trigger Questions, Idea Writing (IW) and Nominal Group Technique (NGT) were applied. The outcome of the Workshop is a list of ranked and organised statements reflecting the implementation phase of the new VLE, addressing positive and negative usability issues.

A three-hour meeting was organised by the authors in April 2018 at the university for academics (n=4), administrators (n=8) and learning technologists (n=1). The participation was voluntary. The aim of the IM session was to collect feedback, discuss questions, problems and capture ideas in connection with the implementation and usability of the new VLE.

3.2.1 Idea Writing

At the beginning of the IM session, the facilitator (one of the author) introduced the methods and the Trigger Questions for the Idea Writing (IW):
Trigger Question 1: What are the **positive** aspects of the implementation of the new VLE?
Trigger Question 2: What are the **negative** aspects of the implementation of the new VLE?

The participants formed two mixed groups (n=6, n=7) and without discussing the question, every participant, focusing on *Trigger Question 1*, wrote one positive aspect of the implementation of the new VLE on his/her paper then passed the A4 sheet to the next member of the group in the circle. After reading the previously listed statements on the new A4 sheet, members wrote another positive statement and circulated the A4 sheets until the original sheets arrived back. The same practice was followed with the *Trigger Question 2*.

3.2.2 Nominal Group Technique

Following the Idea Writing phase, the members of the groups discussed, clarified and edited the positive and negative statements for the preliminary ranking. Each participant selected the five most important statements from the whole list and ranked them by associating numbers from one to five for each statement, five being the most important. Single Transferable Vote technique was utilised to minimise discarded votes during the ranking process.

4. THE RESULTS

4.1 System Usability Scale

4.1.1 The Participants

The quantitative usability evaluation was conducted by utilising the SUS methodology. The total number of participants is n=182 including students (n=137), academics (n=23), learning technologists (n=3) and faculty administrative teams (n=19). Printed (paper) and online questionnaires were offered. N=13 SUS questionnaires arrived on paper evaluated by learning technologists (n=1), academics (n=4) and administrators (n=8). The online questionnaire was submitted by 169 users including learning technologists (n=2), students (n=137), academics (n=19) and administrators (n=11). The evaluation was based on the general experience gained during the first phase of the implementation (from September 2017 to April 2018) of the new VLE by using the features needed for the different user groups. The questions were derived from the original SUS questions (Brooke, 1996) with a slight change in the wording. Unfortunately, a small error slipped into the online student questionnaire. One of the questions was repeated twice and as a result, the last question was left out. This small discrepancy does not affect the result significantly as the structure of the SUS questions and the methodology make the evaluation robust and flexible to small errors and changes (Sauro and Lewis 2011). The standard error is within 0.25 regarding the final SUS score. The accuracy is higher than 99.5 %.

4.1.2 The Interpretation of the SUS Result

The overall SUS score of the new VLE measured after 7 months (April 2018) of the implementation (first phase) is **58.6** out of 100. This is the mean result of the evaluation of n=182 users including students (n=137), administrators (n=19), academics (n=23) and learning technologists (n=3).

There is a well-accepted adjective scale based on the benchmarks set up by Bangor et. al. (2009). A SUS score over 80 suggests a very good, highly usable system, between 68 and 80 is still OK but could be improved, between 51 and 68 means “Fair”, it still works but should be improved, below 51 is poor and below 36 is unusable. The SUS score 68 corresponds to 50% which means that the average score of more than 2300 different systems and products is around 68.

The final score (58.6) is in the range of 51-68 which is below the average usability expectation (68) but still suggest a usable system with scope for improvement.

4.1.3 The Comparison of the User Group Evaluations

A more detailed picture can be seen by analysing and comparing the evaluation of the different user groups. The largest number of users participating in this evaluation are the students (n=137) scored 61.1 opposed to all members of staff (n=45) 49.4. Students’ SUS score weighted more in the overall score and resulted 58.6 for the total average. If the two user groups formed by the students and the staff are weighted equally, the mean SUS score is 55.6, lower than the average score 58.6 calculated with all users as one group. The following chart (Fig. 2) displays the SUS score in respect to the two main user groups, the group average and the total average.

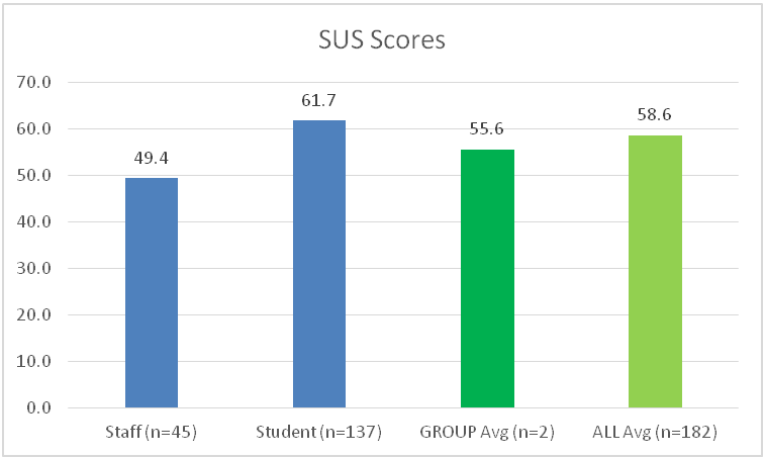

Figure 2. SUS scores of students and staff

4.1.4 Student Group Evaluations

Starting the analysis with the largest user group, the students (n=137), it is interesting to see the comparison of the SUS scores of the different sub-groups within the students.

Student Groups by Levels

Undergraduate (n=127) and postgraduate (n=10) students filled in the online form.

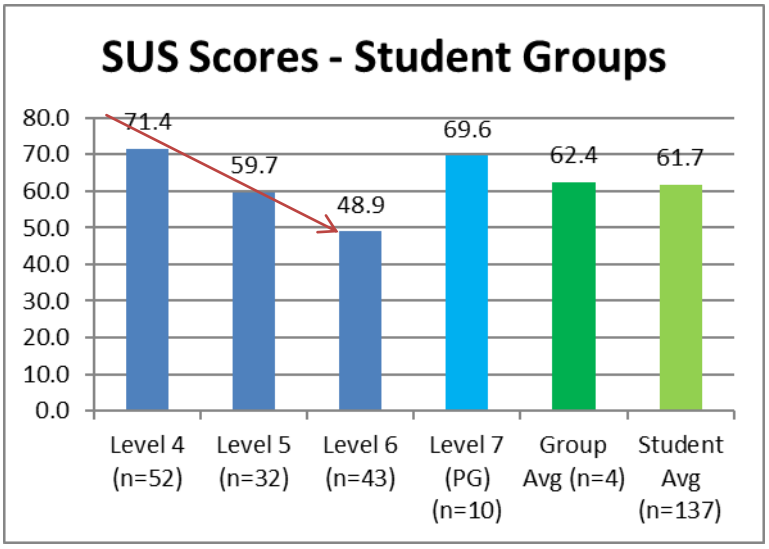

Figure 3. SUS scores of student groups by level

There is a falling trend can which be seen in the graph (Fig. 3) by the undergraduate student groups from 71.4 (Level 4) through 59.7 (Level 5) to 48.9 (Level 6). The first year (Level 4) students evaluated the new VLE slightly above the average expectation. They seem to be satisfied with the new system, unlike the Level 6 students who have higher expectations. The postgraduate students (level 7), however, gave 69.6 for usability (Figure 3) which is near to the generally accepted average (68) for SUS scores.

Student Groups by Frameworks/Courses

The results of six different groups of students can be seen in Figure 4. The groups were formed based on frameworks and courses. The students are from different levels/years in each group. The largest group is the nursing students (n=66). Their average SUS score is 60.3 which is very close to the average score of the six groups (60.2). The difference between the lowest (43.9) and highest SUS score (74.7) is more than 30 (30.8).

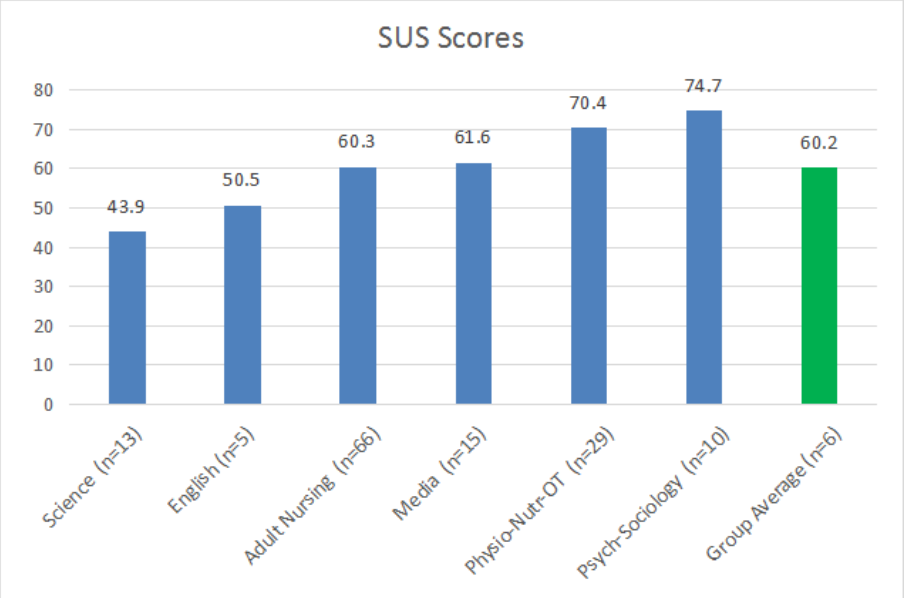

Figure 4. SUS scores of student groups by frameworks/courses

4.1.5 Staff Group Evaluation

N=45 evaluation derived from staff members either online (n=32) or on paper (n=13). The following groups are created: academics (n=23), administrators (n=20), learning technologists (n=3). Figure 5 shows the results graphically. It is conspicuous that academics gave very low usability score (37.8) to the new VLE since administrators and learning technologists scores suggest that the VLE is close to an average system with respect to the usability. The mean value of the groups' SUS scores is 55.4 which is acceptable but the total average falls below 50 (49.4) which is on the borderline of the usability.

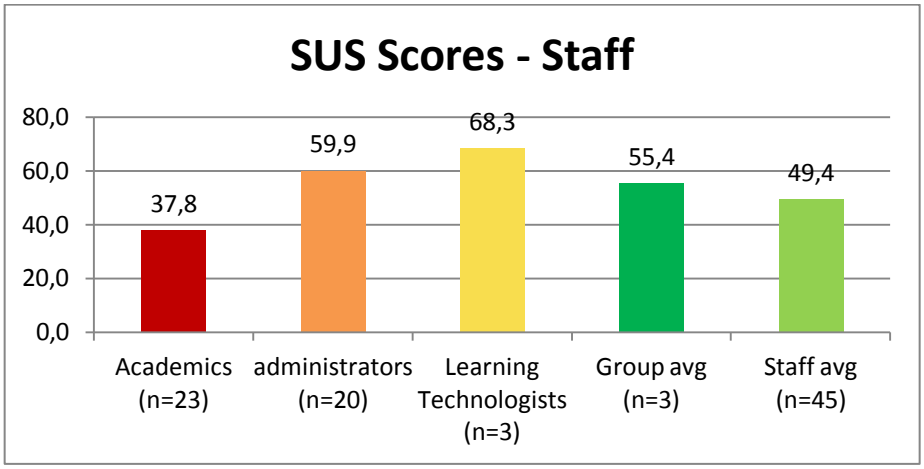

Figure 5. SUS scores of staff groups

The result of the academics (SUS = 37.8) highlights some usability issues. For further analysis, Figure 6 chart shows the individual scores in the academics group (n=23). Blue bars (n=19) shows the online result, yellow bars (n=4) relate to the paper-based questionnaire.

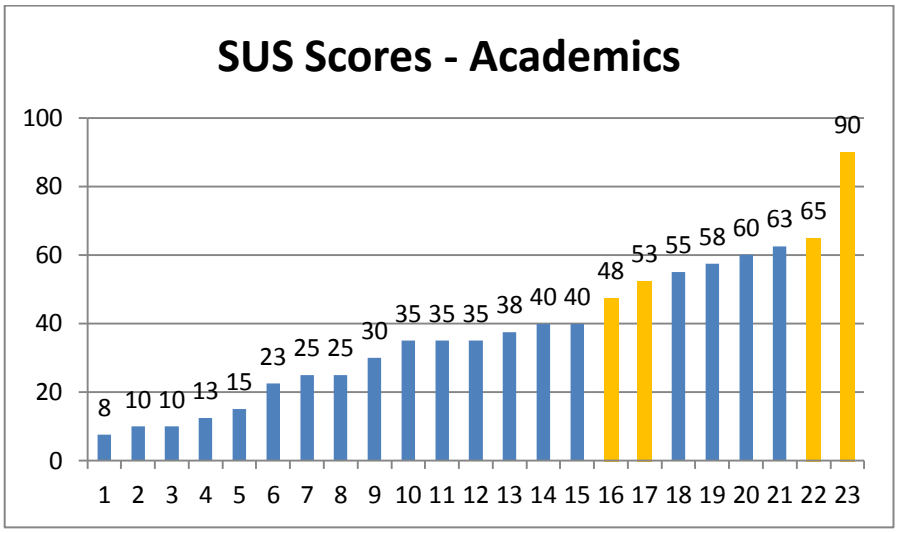

Figure 6. SUS scores of academics

Half of the group of academics (n=12) evaluated the new VLE below 38 which indicate serious usability issues. Interestingly, the paper-based results (n=4) are significantly higher (SUS avg = 64) than the online scores (SUS avg = 32). However, the overall standard deviation is not high (21), the range and distribution of the scores are unusual.

4.1.6 Factor Analysis

A more detailed analysis can reveal the weak areas of the new VLE according to the academics (n=15) who evaluated the system lower than 41. Figure 7 shows the result of each factor (the scores given to each question) of the evaluations which have the total SUS score under 41. These are the first 15 scores from the left on the bar chart in Figure 7.

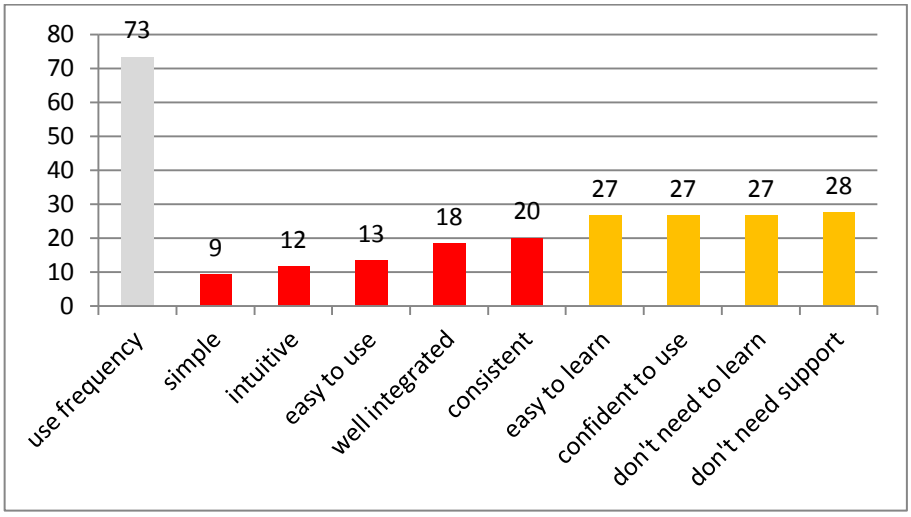

Figure 7. Factor analysis of the weakest evaluations (SUS < 41)

The weakest areas, scored from 9 to 20, are highlighted in red on the bar chart. These academics (n=15 out of 23) did not find the new VLE simple, intuitive, easy to use, well integrated or consistent.

4.2 Interactive Management Session

By the end of the IM session, four lists of ranked statements were produced for the two trigger questions by the two groups. The two positive and two negative lists were merged into one positive and one negative list. The extract of usability related statements are grouped into categories (see Table 1 and Table 2).

Table 1. *Trigger Question 1* Statements

Positive Statements Grouped into Categories
Usability
Clean and fresh, works good, better user interface
Functionalities for staff/students
Programme Support help area now a lot cleaner
Allows students to hand in late submissions in same area, lateness is clearly marked
Learnability
Access to Sandbox to mess around without worrying about breaking the system.
Training organise and run in plenty of time
Advantage in piloting is confidence in year 2
Help section divided for academics/professional support
Support
Able to contact trainers
On-demand help from Learning Tech, contact directly
Programme Support help area now a lot cleaner
Learning technologists were very helpful
Training organised and run in plenty of time
Help section divided for academics/professional support

Table 2. *Trigger Question 2* Statements

Negative Statements Grouped into Categories
Support
No LT support
Training too general
No personal training for unique faculty needs
Too many ways of accessing the same thing
Sandbox can't simulate everything
Usability
Current VLE and new VLE not always linked up
Systems not talking to each other as well as advertised
A lot of things shown were not useful in terms of usability for teaching
No template for structure of unit
Software lacks consistency in interface
Who was consulted regarding large file submission
Communication
Systems not talking to each other as well as advertised
Who was consulted regarding large file submission
Lack of info prior to rolling out
More communication required about Implementation

The positive and negative statements are grouped into categories based on similarities which makes the problem domain clearer and easier to recognise structure and pattern. The order of the statements follows the scores in ranking. The list starts with the most important statements. Some statements are listed in more than one categories if it was required. The categories refer to usability, learnability, support and communication. The individual statements specify the area and nature of the usability issues. IM offers a valuable feedback by supporting the general evaluation of the SUS with specific comments.

5. CONCLUSION

The usability evaluation of the new VLE at this stage provided a reliable and meaningful feedback. The overall SUS score (58.6) suggest a usable system in general but also indicate some usability issues in particular areas. As the implementation is in its early stage (phase 1), this score should not be considered as the SUS score of the fully implemented and fine-tuned system. The analysis of the evaluation of the different user groups and individual users discloses more details and differences within and between the usability perception of the user groups. The VLE is a complex system with numerous features. Each user group evaluates a slightly or significantly different part of the VLE. The divergence between the SUS scores hints that (a) the system is not uniform regarding the usability (b) the expectation and perception are different. The detailed analysis of the low SUS scores (37.8) given by the academics identified five problematic areas: simplicity, intuitiveness, ease of use, integration, consistency. Students are mostly satisfied with the new VLE, although, interesting trends can be seen in the undergraduate results (see Figure 3). Academics and administrators are not always fully satisfied. The IM workshop offered a great opportunity to identify, communicate and resolve some serious usability issues. The feedback captured during the workshop was useful and valuable. The usability evaluation provided a realistic picture of the new VLE at the end of the

first phase of the implementation. The case study offers an example of a feasible usability evaluation of a VLE combining SUS and IM methodologies.

The feedback captured in the IM session give some explanation of the SUS scores. There are more negative statements (n=29) in the ranked lists than positive ones (n=19). The categories refer to the areas that need attention either from the usability perspective or regarding the implementation process. The high importance of support, training, communication is well recognized by the team that manages the implementation and confirmed by the result of this study as well.

The implication of the research:

a) The implementation team gained an overall picture (SUS score) of the usability of the new VLE.
b) The SUS score can be compared to the preliminary and subsequent results.
c) The implementation team could identify usability related issues during the first phase of the implementation and address them at this early stage.

The research has the following limitations. The different user groups were not represented in equal number. Three times more student completed the online evaluation, but no students participated in the IM session. The SUS score comparison of the user groups gives an equal weighting to every user group.

The online SUS questionnaire for students had an error. Question 5 (Fig. 1) was repeated so, question 5 and 6 were the same and the further questions (6-9) were scrolled down to 7 to 10. As a result, the last question (10) has been left out. As the SUS questions were designed with some redundancy for robustness (Bangor et al., 2008), this error did not cause a significant difference in the SUS result of the student evaluation. The phrasing of the first question was modified from 'I think that I would like to use this system frequently' to 'I use this VLE frequently'. The reason behind this change is that there was no choice of using other VLE for these users. The impact of this change is that the SUS scores given to the first question are relatively high compared to the average scores. It slightly raises the mean SUS score.

REFERENCES

Albert, W. and Tullis, T., 2013. *Measuring the user experience: collecting, analyzing, and presenting usability metrics*. Newnes.

Ayad, K. and Rigas, D., 2010. Comparing virtual classroom, game-based learning and storytelling teachings in e-learning. *International Journal of Education and Information Technologies*, *4*(1), pp.15-23.

Babić, S., 2012. Factors that influence academic teacher's acceptance of e-learning technology in blended learning environment. In *E-learning-organizational infrastructure and tools for specific areas*. InTech.

Bangor, A., Kortum, P. and Miller, J., 2009. Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies*, *4*(3), pp.114-123.

Bangor, A., Kortum, P.T. and Miller, J.T., 2008. An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction*, *24*(6), pp.574-594.

Brooke, J., 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry*, *189*(194), pp.4-7.

Broome, B.J. and Keever, D.B., 1986. Facilitating Group Communication: The Interactive Management Approach.

Davis, F.D., Bagozzi, R.P. and Warshaw, P.R., 1989. User acceptance of computer technology: a comparison of two theoretical models. *Management science*, *35*(8), pp.982-1003.

Dogan, H. and Henshaw, M.J.D., 2010. Transition from soft systems to an enterprise knowledge management architecture. In: International Conference on Contemporary Ergonomics and Human Factors 13-15 April 2010 Keele University, UK.

Emelyanova, N. and Voronina, E., 2014. Introducing a learning management system at a Russian university: Students' and teachers' perceptions. *The International Review of Research in Open and Distributed Learning*, *15*(1).

Gallud Lazaro, J.A., Albertos Marco, F. and Ruiz Penichet, V.M., 2013. Collaborative e-learning through drag & share in synchronous shared workspaces.

Granic, A. and Cukusic, M., 2011. Usability testing and expert inspections complemented by educational evaluation: A case study of an e-learning platform. *Journal of Educational Technology & Society*, *14*(2), p.107.

Lewis, J.R., 1991. Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the ASQ. *ACM Sigchi Bulletin*, *23*(1), pp.78-81.

Lewis, J.R., 1995. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, *7*(1), pp.57-78.

Luo, G.H., Liu, E.Z.F., Kuo, H.W. and Yuan, S.M., 2014. Design and implementation of a simulation-based learning system for international trade. *The International Review of Research in Open and Distributed Learning*, *15*(1).

Melis, E., Weber, M. and Andrès, E., 2003. Lessons for (pedagogic) usability of eLearning systems. In *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* (pp. 281-284). Association for the Advancement of Computing in Education (AACE).

Nielsen, J. and Phillips, V.L., 1993, May. Estimating the relative usability of two interfaces: heuristic, formal, and empirical methods compared. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems* (pp. 214-221). ACM.

Orfanou, K., Tselios, N. and Katsanos, C., 2015. Perceived usability evaluation of learning management systems: Empirical evaluation of the System Usability Scale. *The International Review of Research in Open and Distributed Learning*, *16*(2).

Paulsen, M.F., 2002. Online education systems: Discussion and definition of terms. *NKI Distance Education*, *202*.

Renaut, C., Batier, C., Flory, L. and Heyde, M., 2006. Improving web site usability for a better e-learning experience. *Current developments in technology-assisted education*, pp.891-896.

Ryan, S., Scott, B., Freeman, H. and Patel, D., 2013. *The virtual university: The internet and resource-based learning*. Routledge.

Tullis, T.S. and Stetson, J.N., 2004, June. A comparison of questionnaires for assessing website usability. In *Usability professional association conference* (Vol. 1). (pp. 7–11).

Venkatesh, V. and Davis, F.D., 2000. A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management science*, *46*(2), pp.186-204.

Warfield, J.N. and Cárdenas, A.R., 1994. *A handbook of interactive management* (p. 338). Ames: Iowa State University Press.